



Scars collaborative telediagnosis platform using adaptive image flow

Rami Kassab, Jean-Christophe Lapayre, Jean-Baptiste Aupet, Franck Marzani, Christian Pieralli

► To cite this version:

Rami Kassab, Jean-Christophe Lapayre, Jean-Baptiste Aupet, Franck Marzani, Christian Pieralli. Scars collaborative telediagnosis platform using adaptive image flow. Integrated Computer-Aided Engineering, 2013, 20 (1), pp.3-14. hal-00790426

HAL Id: hal-00790426

<https://hal.science/hal-00790426>

Submitted on 20 Feb 2013

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Scars collaborative telediagnosis platform using adaptive image flow

R. Kassab^a, J.-C. Lapayre^{a,*}, J.-B. Aupet^a, F. Marzani^b and C. Pieralli^a

^a*FEMTO-ST Institute – UMR CNRS, Franche-Comte University, Besancon Cedex, France*

^b*Le2i Laboratory Bourgogne University – UMR CNRS, Dijon Cedex, France*

Abstract. Telemedicine has been developed to allow practitioners to remotely connect with patients and with other medical staff. We propose a new system (hardware and software), named DICODERM (COLlaborative DIagnosis of DERMatosis), which makes it possible to monitor the evolution of scars after the excision of a tumorous dermatosis (like melanoma). The hardware part of this system is composed of a new optical innovative probe with which two types of images can be acquired simultaneously: anatomic with a white light image and functional with a fluorescence image (using autofluorescence from the protoporphyrin within the cancer cell).

The software part is composed of two components: the image stitching component, and the collaborative/adaptive layer component. Our system creates a panoramic view of these scars obtained by stitching a sequence of small images. We conducted experiments for different image stitching algorithms to define the best solution.

We also deployed a second component: a collaborative system layer which allows to remotely share images of scars and to adapt these images. We also made the system adaptive to communicate across different client platforms. We conducted experiments to compare the exchange of images with or without adaptation: these tests showed the efficiency of our layer.

Keywords: Adaptability, collaborative medical telediagnosis, stitching, telemedicine, scar monitoring

1. Introduction

When a tumorous dermatosis (like a melanoma) is detected early enough, the prognosis is favorable after excision. Subsequently, the result of scar excision should be monitored. However, in addition to cicatrization, it is necessary to check whether all cancer cells have been removed. The vital prognosis of the patient being life-threatening, it is interesting to involve different remote practitioners to promote collaborative work.

In the project DICODERM (COLlaborative DIagnosis of DERMatosis), we propose a new complete system (hardware and software) for collaborative monitoring of scars.

The first part is the hardware. In our laboratory CNRS Femto-ST, we have developed a new optical probe that can acquire images in two modes (white light and fluorescence). The white light image is a classic image of the scar: known anatomical image. The fluorescence image is obtained by showing the autofluorescence of protoporphyrin content in the cancer cell. The originality of this system is to allow the simultaneous acquisition (by the same system) of the two types of pictures.

The acquisition system only produces macro images of a small area of skin. We study different stitching algorithms to determine the most appropriate ones to reconstruct the panorama of the whole scar. Depending on the particularity of these images, we set up a series of tests to choose and to develop our component stitching. As our system has a single source of acquisition, the system can use the registration parameters of conventional images to affect the registration of fluorescence images.

*Corresponding author: Jean-Christophe Lapayre, FEMTO-ST Institute – UMR CNRS, Franche-Comte University, 16 route de Gray, F-25030 Besancon Cedex, France. E-mail: jean-christophe.lapayre@femto-st.fr.

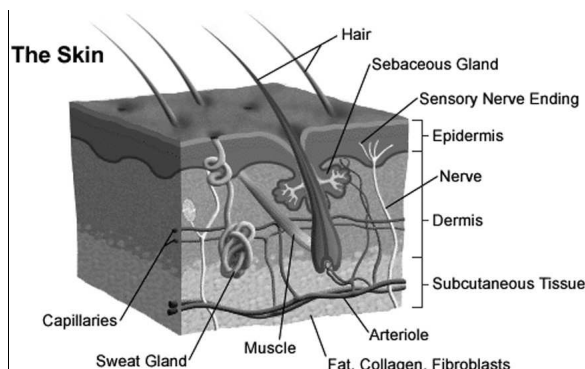


Fig. 1. Diagram of skin composition.

Telediagnosis is the process of connecting several practitioners via a robust communication system. It saves both time and money and provides the opportunity of sharing the different experiences of different medical staffs. This kind of application improves the chances for the patient's medical condition to improve.

The system then shares the two panoramic images (white light and fluorescence) in a robust and reliable display: so practitioners can work together in a collaborative telediagnosis. Due to the different types of terminals and different types of networks, our platform should allow the adaptation of data.

In the first section, we present the contexts (medical, image treatment and collaborative telediagnosis) of this research work. In the following section, we describe our contribution to providing a complete integrated system, from initial acquisition and treatment of images to their use in collaborative telediagnosis. The last section gives details about implementation and the performance results. The last section presents the conclusion and further work.

2. Context

Over the past 10 years, telemedicine and telediagnosis have grown considerably. This phenomenon is due in large part to the parallel increase of performance networks and processors, but also due to the improvement of security in these systems.

Teleworking is used in various ways such as distant learning, remote maintenance and even telemedicine or teledermatology [14,36].

This research work is based on several domains of research: the optic domain with the use of an optical probe to prominently display the autofluorescence of protoporphyrin within the cancer cell; the image pro-

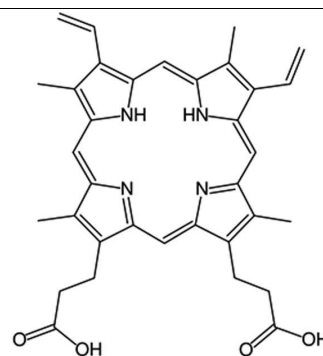


Fig. 2. Molecular structure of PPIX.

cessing domain for stitching the acquired images; and the collaborative work platform domain for sharing and adapting data exchanges. That is why this section is divided into three sub-sections.

2.1. Medical context: System cancer cell detection in skin

Skin tumors occur mainly when the human body is overexposed to sun or due to genetic disorders. Mainly, two kinds of cancers can be cited: melanoma and basal cell carcinoma. The prognosis is not at all the same and the screening and monitoring of scars after excision is of critical importance. Our tool is based on fluorescence. Indeed, a protein, protoporphyrin IX (PPIX) is present in all skin cells. It is produced by the cells and is then consumed and finally disappears; this is called the deterioration cycle of PPIX. Nevertheless, this cycle is highly disrupted in case of malignant cells, and even disappears. That is why in this case, there is a large concentration of PPIX in malignant cells. Because PPIX fluorescence is well known, measuring the cell auto fluorescence is therefore a powerful tool in skin cancer screening.

The acquisition device is based on one property of the skin. For example, we present in Fig. 1 a diagram of skin composition.

The skin is composed of two layers: the epidermis and the dermis. They are made of several cell types (e.g. melanocytes). All cells produce different proteins. The proteins are created, used and then destroyed for creating energy or for other reasons. The problem is when some cells are malignant. Two types of consequences can occur: proteins are no longer synthesized or, on the contrary, proteins are no longer consumed. It appears that a specific protein, protoporphyrin IX, noted PPIX in Fig. 2, follows the second pattern pre-

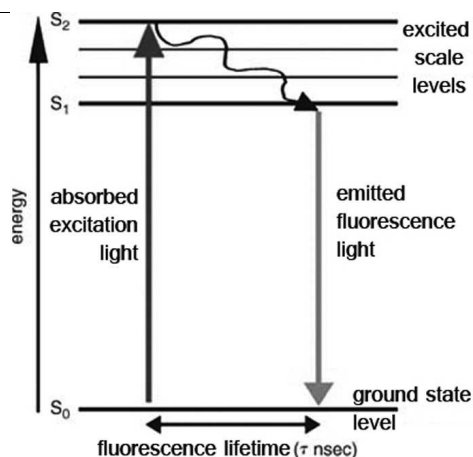


Fig. 3. Jablonski diagram.

sented above. That is to say that this protein is always synthesized but is not degraded by the cancer cells. As a result, this protein accumulates in the malignant cells.

As we can see in Fig. 2, PPIX is composed of several kinds of chemical connections, σ , Π and nitrogenated pentacycles. The latter exhibit an interesting property, since they can induce fluorescence. This phenomenon can be easily described by the Jablonski diagram Fig. 3.

As we can see, it represents the energy state levels of PPIX molecules. The phenomenon is due to absorbed light excitation by the molecules. Electrons, originally at the ground state level S_0 absorb photons and rise to excite the state level S_2 . In the next step, electrons fall down to state level S_1 by emitting heat. Finally, they fall back down to level S_0 . During this stage, there is an emission of light and this is called fluorescence. As we can see, the photon absorption energy gap (S_0 to S_2) is larger than that of the emitted light (S_1 to S_0). In other words, the emitted light wavelength is larger than that of light excitation. It is this process that we use to determine the presence of malignant cells. Indeed, as the concentration of PPIX is higher in cancer cells than in healthy ones, if we adequately illuminate the skin, we can distinguish the investigated cells which are abnormal (high intensity of fluorescence) from those which are normal.

It appears that in the case of PPIX, the absorption peak is at a wavelength of 406 nm for a fluorescence one centered around 670 nm. It must be mentioned that fluorescence exhibits another property that we used to create a device devoted to skin cancers in order to demonstrate the feasibility of our method. If we illuminate the molecules, not with a continuous light [5] but rather with a pulsed light (succession of flashes),

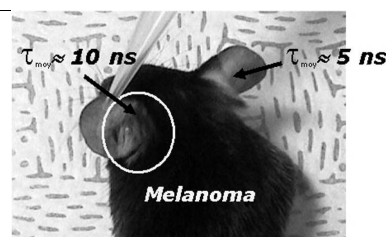


Fig. 4. Mouse model: LT of cancerous ear (left in the picture) is 10 ns, 5 ns for the normal one (right).

the radiative behavior of fluorescence emission evolves temporally. There are two steps: the molecules absorb light during the flash and then re-emit fluorescence for a certain time. This is called Fluorescence Lifetime (FL). It appears that FL clearly depends on the molecule environment. We used thus this property to investigate melanomas with mice by inoculating cancer cells in one ear. As we can see in Fig. 4, the PPIX fluorescence lifetime of the normal ear is 5 ns whereas that of the cancerous ear is 10.

This result shows the pertinence of using fluorescence to investigate cancers. The device was then abandoned because of its limitations: point measurement (no pictures), experimental difficulties and cost. Therefore, we created a device also based on fluorescence but for its intensity rather than its lifetime, as it will be shown later (Section 3.1).

In the state of the art, different fiber sensors are proposed for detecting fluorescence in the skin. Previous systems used either one fiber to transmit the blue light and to receive the fluorescence reflected by the skin when the distance between the sensor and the skin was not sufficient or a fiber bundle to transmit and receive when the distance was adequate but spatial resolution was determined.

In Fig. 5 we can see the hybrid system [10] which uses one fiber to transmit the blue light and four or more fibers to receive the fluorescence reflected by the skin. The hybrid system allows a large distance of working space while maintaining a good spatial resolution.

The system we need has to consider the optical issue in its functionality; our acquisition system is explained in the contribution section.

2.2. Image registration and panoramic image stitching

2.2.1. Introduction

Image registration is the process of overlaying two or more images of the same scene taken at different

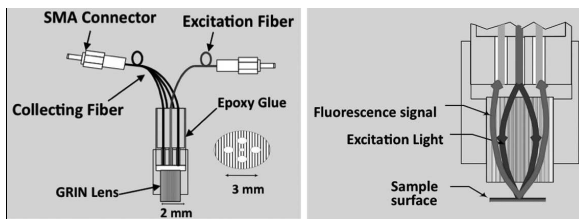


Fig. 5. Previous optical probe acquisition system.

times or from different viewpoints, and/or by different sensors. It geometrically aligns two images: the reference and sensed images [28].

Image registration techniques depend on the type of acquisition system: in the case of a multi-view analysis, the goal is generally to reach a 3D representation of the scanned scene [26,32]. When pictures of the same scene are acquired at different times, we use the word multi-temporal. In this case, the aim is to find and to evaluate changes in the scene, which appears between the consecutive image acquisitions.

Sometimes, in medical imaging monitoring of the healing process, or in monitoring tumor evolution, different sensors (with multimodal analysis) are used and images of the same scene are acquired by various sensors [22,27,37].

Panoramic image stitching has several research domains and commercial applications. Nevertheless, the first results in this domain did not consider the approach of automatic stitching, because they needed human input or restrictions on the image sequence for establishing matched images. In 2007, David Lowe proposed a solution based on his algorithm called SIFT for *Scale-Invariant Feature Transform* [17,18], giving an invariant feature based approach to fully automatic panoramic image stitching [7]. His method is insensitive to the ordering, orientation, scale and illumination of the input pictures. It is also insensitive to noisy images that are not part of a panorama, and can recognize multiple panoramas in an unordered image dataset. There has been other research on automatic panorama by detecting key points [27] and automatically matching them [1,9,11,33]. Other studies have been conducted concerning mobile applications where the user can take an image sequence for a wide range of scenes with a camera phone, and see a panoramic image of the scenes on the camera phone immediately [38]. Still other research has tried to find new algorithms to enhance performance by decreasing processing time [24,30].

Among the different existing image stitching algorithms dependent on the description of all state of the

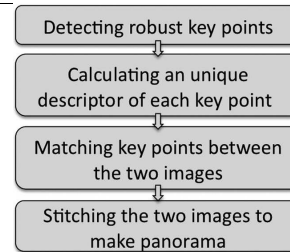


Fig. 6. Stitching steps.

art algorithms, we found that SIFT, ASIFT [25], and SURF were the more suited to our work. We work with skin cancer images of two types: “white light and fluorescence”. We applied our implementations and tested these three algorithms to obtain comparable results and to choose the best scenario for our work. As proved in the contribution paragraph, SIFT and ASIFT are more robust than SURF, but SURF is faster than ASIFT and SIFT because it is applied on integral images [4].

2.2.2. Image registration algorithms

In this paragraph we give details of the three algorithms briefly described in the previous paragraph, to highlight the different techniques used in each of them. Generally, most image stitching methods (methods on a pair of images) follow essential steps as in Fig. 6 where detecting characteristic features in images is the first step. Usually these features will be detected in the next successive image and help in applying the following step which finds the match between the two images. Depending on the match, we can find the transforming parameters to register images and stitch them together.

a. SIFT algorithm

Detecting key points was primarily designed for robotic applications in order to recognize objects within a captured image using a visual database as a reference, but Lowe and Brown [7] also showed a practical application of image registration for automatic panorama recognition.

The three major stages of computation used to generate the panorama in the SIFT algorithm [22] are:

1. Scale-space extrema “key-point” detection: the first stage of computation seeks these points over all scales and image locations. It is efficiently implemented by using a difference-of-Gaussian function to identify potential interest points that are invariant to scale and orientation.

$$\begin{aligned}
D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) \\
&\quad * I(x, y) \\
&= L(x, y, k\sigma) - L(x, y, \sigma) \quad (1)
\end{aligned}$$

where

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (2)$$

$L(x, y, \sigma)$ is produced from the convolution of a variable-scale Gaussian, $G(x, y, \sigma)$, with an input image, $I(x, y)$. The parameter σ refers to the smooth parameter of Gaussian function, and the difference-of-Gaussian function convolved with the image, $D(x, y, \sigma)$, which can be computed from the difference of two nearby scales separated by a constant multiplicative factor k .

2. Key-point localization: at each candidate location, a detailed model is fit to determine location and scale. Key-points are selected based on measures of their stability.
3. Orientation assignment: one or more orientations are assigned to each key point location based on local image gradient directions. All future operations are performed on image data that has been transformed relative to the assigned orientation, scale, and location for each feature, thereby providing invariance to these transformations.

$$\begin{aligned}
m(x, y) &= \\
&\sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (3)
\end{aligned}$$

$$\begin{aligned}
\theta(x, y) &= \tan^{-1}(L(x, y+1) - L(x, y-1)) / \\
&\quad (L(x+1, y) - L(x-1, y)) \quad (4)
\end{aligned}$$

4. Key-point descriptor: local image gradients are measured at the selected scale in the region around each key-point. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination [17].

b. SURF algorithm

The approach for interest point detection uses a very basic Hessian-matrix approximation, because of its reliable accuracy. More precisely, blob-like structures at locations where the determinant is optimum are detected. In contrast to the Hessian-Laplace detector by

Mikolajczyk and Schmid [24], the SURF algorithm also relies on the Hessian determinant for scale selection, as performed by Lindeberg [15].

The major steps of image feature detection of SURF and SIFT are similar, but SURF lends itself to the use of integral images as made popular by Viola and Jones [34], thus reducing the computational time drastically. Integral images fit in the more general framework of boxlets, as proposed by Simard et al. [31]. The major goal of using integral images in SURF is to decrease the time of computing key-points. They allow for fast computation of box type convolution filters.

An integral image, denoted $i(x, y)$, at location (x, y) contains the sum of the pixel values above and to the left of (x, y) . Formally:

$$i(x, y) = \sum_{x', y' \leq x, y} (i(x', y')) \quad (5)$$

c. ASIFT algorithm

This algorithm depends essentially on SIFT and proposes a crucial parameter for evaluating the performance of affine recognition, the transition tilt. The transition tilt measures the degree of viewpoint change from one view to another, giving a first intuitive approach to absolute tilt and transition tilt. It illustrates why simulating large tilts on both compared images proves necessary to obtain a fully affine invariant recognition. Indeed, transition tilts can be much larger than absolute tilts. In fact they can behave like the square of absolute tilts [25].

2.3. Collaborative virtual environments (CVE) for tele-medicine

A collaborative workspace or shared workspace is an inter-connected environment in which all the remote participants can access and interact with each other just as inside a single entity.

The environment may be supported by electronic communications and groupware, which enable participants to overcome space and time differentials. This is typically made possible by a shared mental model, common information, and a shared understanding by all of the participants regardless of physical location.

In order to become more innovative and competitive, companies choose software for new forms of collaboration. Instant messaging allows direct communication and collaboration in real time. It integrates applications and existing business processes and allows collaboration seen in the context of the task at hand. The collabor-

orative tools are used by communities to provide team members with instant access to secure content sharing and the expertise of colleagues who may be nearby or across the globe [23]. Collaboration tools and awareness features are used to make new software more efficient than standard software in terms of collaborative work level.

2.3.1. Medical telediagnosis

Telemedicine and e-Health are the use of electronic communication technologies as a method of providing health care, education, and related services (medical imaging, distance teaching, patient file consultation). Dissolving barriers such as distance, time, and geography, these applications are designed to improve access to specialized health care.

Telemedicine is generally used in a non-acute setting for patient monitoring or education [17] and has only recently been introduced into emergency care [20, 21]. Telemedicine can be defined as the use of telecommunication technologies to provide medical information and services [19]. In particular, medical telediagnosis is the process by which electronic, visual and audio communications are used to help practitioners in different places to perform diagnostic procedures and consultations, such as clinical examinations through the transfer of medical images. These images are often large and may take time to be routed according to the networks used.

The goal of telemedicine is to allow practitioners to act as if they were in the same room, using a wide panel of medical tools [8]: like a virtual examination room. With these applications, practitioners can exchange their information and experience as if they were in the same room working together. Collaborative Virtual Environments (CVE) provide all the services to perform these applications. A telemedicine application is collaborative but with special constraints: the patient's life is at risk. So, this kind of environment should guarantee communication, security, sharing management, and concomitant access management.

The users of these applications are mobile, and they work on different terminals, on different networks, ...: heterogeneity is omnipresent. Thus adaptability is needed.

2.3.2. Adaptability

A lot of criteria could change the system to accommodate user waiting time, but users will not waste their time manually configuring system properties, especially if they want to use the system with different

terminals: for example with their own computer or a public access point with a PDA (Personal Digital Assistant) [29,39].

If users have to change the configuration of their system each time they change work environments, they quickly abandon this system to avoid wasting time. Therefore the system should be adaptive, and provide an environment that evolves as use evolves. That is the reason why we introduce the notion of adaptive systems. The difference between adaptive and adaptable refers to the extent to which users can influence the individualization process of a system. Adaptive systems are customized by the users themselves, whereas with adaptable systems this process will be done automatically [16].

The ultimate goal of a collaborative system is to be able to adapt to changing initial needs, to emerging needs and to changes in contexts and technologies. The term co-evolution [39] was chosen to reflect the fact that collaborative systems must evolve continuously, but not with total autonomy or self-adaptation, because they should take into account changing needs, changing attitudes and the changing skills of users, individually or collectively. It then becomes necessary to dynamically adapt an application itself and the data which transit between nodes according to the characteristics of the executive environment [13,14].

3. Contributions

3.1. Image acquisition

As mentioned above, applying image processing such as rotation, scaling, or panorama to conventional images acquired under white light is much easier than applying it to fluorescence images. That is why we use the acquisition system depicted in Fig. 7, which enables us to take two images of the same position, once under a white light and once in fluorescence mode.

In order to obtain an accurate picture of a scar in fluorescence and white light Fig. 7, there is a system that uses an optical probe with a CCD camera on the same axis for white light and blue light that gives us two images describing exactly the same area.

The probe we use has a special size limited to the size of blue light, so it cannot cover the whole area of a scar or cancer. Therefore, we have a sequence of small images of the requested area of skin.

Hence we can calculate the registration parameters of conventional images acquired under white light.

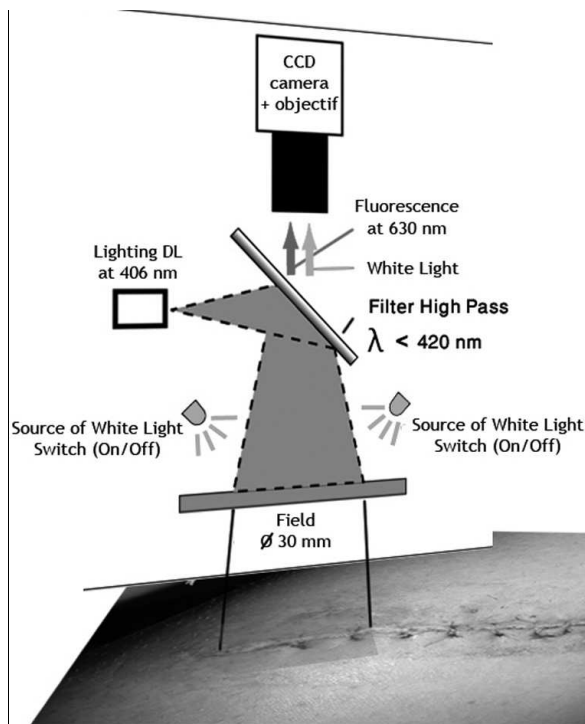


Fig. 7. Image acquisition process.

Registration parameters describe the transformation (rotation, translation, scale ...) between each pair of images. These parameters are used to achieve the whole panoramic fluorescence image.

The acquisition of a set of image pairs (one under white light and another in fluorescence mode) provides information on the whole area of skin. This means that we will be able to process the registration parameters from conventional images to build a panorama from the set of fluorescence images Fig. 8. It will be possible if a common area of skin always appears between each pair of successive images.

3.2. Global platform

Our local platform is divided into two main parts, which have been requested by medical practitioners:

1. Multi-view,
2. Multi-temporal.

In multi-view tab the practitioner can import pictures (white light and fluorescence) and can ask the system to recover the panoramic image of both conventional and fluorescence images. In multi-temporal tab, the system has different panoramas of the same scar taken at different times. The purpose is to compare

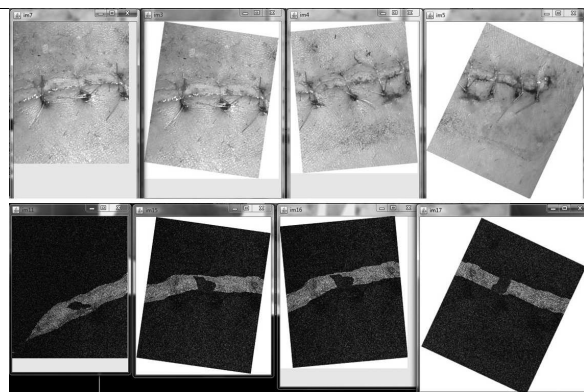


Fig. 8. Image acquisition results.

these panoramas and to observe how the scar is healing. The mission of the system here is to register the panoramas making them comparable for practitioners.

As previously mentioned, the image-stitching process consists of several sub-processes. In our case, we suppose that the changes between successive pictures could be limited to rotation, scaling and sometimes perspective. Our system is a component-based one. So, for example, if a practitioner needs a treatment like lightening effects or 3D effects, it will be easy to integrate new components.

The stitching algorithm, applied to two consecutive images, allows our system to be able to define the transformation parameters between source (first image) and target (second image) images. The system then detects the overlapping area between them, and transforms and adapts the source image to the target one. Stitching is done to obtain a panorama of these two images. This panorama is considered as the first image in next step. We repeat the first step with other images until the final panoramic image is obtained by concatenation of the targets.

We can see from the fluorescence images Fig. 8 that it is difficult to detect any extrema [17]. Hence we cannot calculate the transformation parameters from these images. In addition to the white light and fluorescence pictures which have been taken from the same sources (this is one of our system's features, both innovative and original), transformation parameters of the conventional images are the same and can be applied on the fluorescence images. Then we finally succeed in stitching these images to create the panoramic fluorescence image.

3.3. Stitching algorithm

Each image registration process has three essential steps: key-point detection, key-point matching and

transformation.

The information estimated by the key-point matching step is used to calculate 2D homography. In the context of computer vision, homography almost always refers to mapping between points on two image planes that correspond to the same location on a planar object in the real world. It can be shown that such a mapping can be represented by a single 3-by-3 orthogonal matrix [6]:

$$H = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \quad (6)$$

The elements of this matrix represent different types of transformation: a_{11}, a_{12}, a_{21} and a_{22} are responsible for rotation and scaling; a_{13}, a_{23} are responsible for translation; and a_{31}, a_{32} are responsible for projective transformation where a_{33} is usually equal to 1 when we divide by a_{33} to obtain the 2D coordinates from the 3D ones.

For example point A' from the second image could be changed by all the above-mentioned transformations by multiplying with H to obtain point A in the resulting image.

$$\text{Let } A = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}, A' = \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$$

$$A = H.A' \quad (7)$$

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & 1 \end{bmatrix} \cdot \begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}$$

$$x = \frac{a_{11}.x' + a_{12}.y' + a_{13}}{a_{31}.x' + a_{32}.y' + 1} \quad (8)$$

$$y = \frac{a_{21}.x' + a_{22}.y' + a_{23}}{a_{31}.x' + a_{32}.y' + 1} \quad (9)$$

Therefore, the use of homography in our work transforms the second image by multiplying every pixel inside it by the H matrix. We then obtain a new image aligned to the first one which is the essential step in image stitching. H elements have been estimated by relating matched points between the two images. We can conclude that the more robust matching points we have, the better homography we can estimate and then the better panorama we obtain. To achieve robust matching, there are many algorithms which differ in terms of speed and invariance for different types of transformation.

We tried to mix the three solutions which take advantage of the three algorithms (SIFT, ASIFT, SURF) simultaneously.

In previous studies, we used Harris algorithms to detect the key-points. That is why we had HFVD [35] as a descriptor to be adapted with Harris key-points. After preliminary studies, we decided that the Harris detector was not efficient enough for our type of pictures and we looked for other detectors and descriptors. Finally, we chose between SIFT, ASIFT, and SURF.

With regard to the state of the art, we find that the SIFT descriptor is more invariant toward different changes than the SURF one but slower, and at the same time it is less invariant than the ASIFT descriptor but faster. This is why we decided to build our application essentially with the SIFT descriptor in most of our case studies because it is conveniently robust and fast.

Keeping in mind that in some cases SIFT could make some errors, using SIFT alone would not ensure the robustness of our system. That is why we propose the next scenario.

If high speed is essential, we start by using SURF. If the result is not good enough “not enough matching points to make a panorama” we repeat the work with SIFT. If this is still not good enough we use ASIFT.

If speed is important but not an essential condition, we start with SIFT. If the result is not good enough, we apply ASIFT. If speed is not important, we start with ASIFT.

In such a way, we can guarantee the robustness for most of the examples in our case study.

In order to apply this scenario, two types of tests are required, the pretest that makes us able to predict the situation and the limits between the algorithms, and the test at the end of the panoramic process where we check the accuracy of the new position of pixels after transformation.

3.4. Scar image stitching for remote collaborative diagnosis

In our case we can take the example of optical probes, manipulated by hospital staff in Quebec and practitioners located in other different towns in Canada (for example Ottawa and Montreal). To make this system work well we need an efficient collaborative system that guarantees data sharing in robust and secure ways.

The scenario will be as follows:

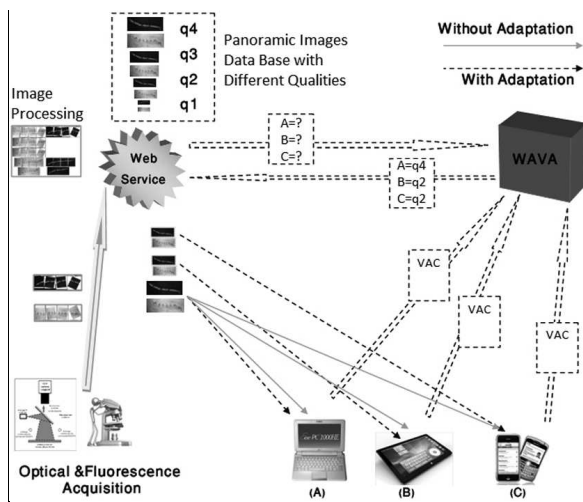


Fig. 9. Collaborative system for stitching images co-interpretation.

1. Hospital staff take the conventional and fluorescence images, and send them to a web service by any possible connection (internet, mobile, satellite, ...).
2. The web service receives these images and manipulates them producing the requested panorama of both conventional and fluorescence modes.

The web service then distributes these two panoramas to all practitioners who are involved in the telediagnosis. In Fig. 9, we can see that the system sends the same panorama to each practitioner.

There could be a problem with the quality of service, but this was our first step in collaborative work without using any adaptation process. In the next paragraph we present our solution for adapting data distribution.

3.5. Adaptability collaboration enhancer

We suppose that the practitioners could have different types of connections to the system (for example Internet, Wi-Fi, 3G, ...), and as well different terminals with different capabilities. Because they cannot receive the same panorama, we switch to the adapted solution by adding the WAVA “*new Web service for Automatic Video data flow Adaptation*” module, which adapts the previous scenario like this:

1. Practitioners (A, B, C) connect to WAVA to register themselves, transmitting their capabilities as VAC [2,3], WAVA will give each of them a rank.
2. The web service receives the images from the acquisition component. It applies the process to produce both the panoramic conventional and fluorescence images, making different copies with different qualities, saving them in a database.

3. When practitioners connect to the web service searching for panoramic images, the web service sends a quality request to WAVA, which sends back the adapted data. At the end, the web service sends to each practitioner the adapted panorama, from which he will make his diagnosis (dotted gray arrows in Fig. 9).

4. Implementation and tests

Our test implementation is divided into two major sub-tests. The first test is for the image stitching step where we compare the results of different existing stitching algorithms applied over our images, through different scenarios. The second test compares the performances obtained by our adaptive collaborative system and other collaborative systems.

4.1. Image stitching implementation and tests

In this case, efficient tests are required to obtain a robust implementation. These tests are supposed to be applied over various image processing algorithms. We then calculate the estimated time for the process of full automatic stitching in both modes: white light and fluorescence. Finally, we share this panorama using our collaborative system mentioned in the previous paragraph Fig. 9.

First, we decided to make our own pretest to have comparable values with results.

In order to do that, we start our test from existing panoramas which we then cut into 4 or 5 smaller images. Then we apply transformations with specific rotation, scaling, and perspective parameters over these sub-images. In the end, we have a table of transformation values which have to be close to the values obtained with SIFT, ASIFT, and SURF. We call this step the pretest.

The next step is to apply the methods we mentioned above, and to get the results. We then compare the results from these three algorithms to the pretest.

The last step is to compare the pixels transformed in the second image to their original positions in the original image.

Table 1 shows the difference, in terms of performance, between SIFT, ASIFT and SURF depending on calculation time and the number of matching points detected between the images. The higher the number of matching points, the more accurate the method should be.

Table 1

Time and number of matching for each algorithm

| Pre-transformation | | SIFT | | ASIFT | | SURF | |
|--------------------|-----------------|----------|----------|----------|----------|----------|----------|
| Overlap % | Rotation degree | Time sec | Match nb | Time sec | Match nb | Time sec | Match nb |
| 30 | 10 | 3.336 | 45 | 196 | 184 | 0.171 | 11 |
| 30 | 15 | 2.391 | 32 | 185 | 126 | 0.162 | 8 |
| 30 | 20 | 2.333 | 31 | 197 | 100 | 0.164 | 4 |
| 30 | 30 | 2.872 | 20 | 184 | 55 | 0 | 0 |
| 20 | 10 | 2.6291 | 23 | 182 | 86 | 0 | 0 |
| 20 | 15 | 2.730 | 17 | 170 | 43 | 0 | 0 |
| 20 | 20 | 2.433 | 13 | 182 | 27 | 0 | 0 |
| 20 | 30 | 0 | 6 | 200 | 21 | 0 | 0 |
| 10 | 10 | 0 | 7 | 0 | 0 | 0 | 0 |
| 10 | 15 | 0 | 4 | 0 | 0 | 0 | 0 |
| 10 | 20 | 0 | 0 | 0 | 0 | 0 | 0 |
| 10 | 30 | 0 | 0 | 0 | 0 | 0 | 0 |

Table 2

Time and number of matching for each algorithm with perspective transformation (extract of results)

| Pre-transformation | | | SIFT | | ASIFT | | SURF | |
|--------------------|-----------------|-----------------|----------|-------------------|----------|----------|----------|----------|
| Over- lap % | Rotation degree | Perspec- tive % | Time sec | Match nb | Time sec | Match nb | Time sec | Match nb |
| 30 | 10 | 0.005 | 2.182 | 22 with errors | 169 | 124 | 0 | 0 |

with errors

To include most of the cases that we might deal with in clinical tests, we prepared various conditions of images (Table 1). In this table, we have a sequence of images with overlapping values varying from 10% to 30% and rotation angles varying from 10 degrees to 30 degrees. In this way, our results confirm that with large overlapping and a small rotation angle, the number of matching points is higher and the obtained panorama is more robust. (The applicable results are framed in bold in Table 1).

Table 2 shows that with 30% of overlapping and 10 degrees of rotation, if we just apply a perspective transformation of 0.05%, ASIFT is the only one which obtains a good result.

The pretest we propose could be applied on images for which we already have panoramas, and the goal is to evaluate our methods. With real images we will apply another type of test which will depend on the result itself without any other known parameters like in the pretest.

Our implementation was first programmed in Java, and then we imported our code to a web service that uses a Tomcat server and connects to MySQL database where we stored our panoramas with different resolutions. Then to make efficient tests on our images, we

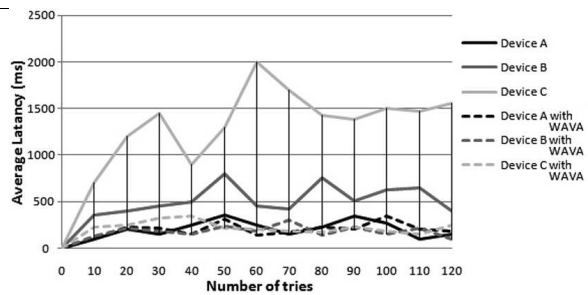


Fig. 10. Adapted collaboration results.

use Open CV “Open Computer Vision” library in C language, where many procedures already exist to perform the desired work. At the level of image processing, our platform is designed to test the efficiency of work distribution for the algorithms that we use in the image stitching step.

4.2. Collaboration implementation and test

Collaborative applications allow users to share data, but if terminals and networks are heterogeneous it is hard to guarantee that no data will be lost. In these tests, we show the quality of our system which is able to adapt data and then provide a system with no loss. Figure 10 compares our system and other collaborative systems that do not integrate any adaptation. These results show that with no adaptation we lose a substantial number of packets with B and C which have only a small capacity (CPU and network) like that of cell phones or PC-tablets. When adapting the panorama resolution to the matched capacity of each practitioner’s device, the number of lost packets is significantly reduced. In this test we consider the process of stitching images as a black box which is performed on the web service.

5. Conclusion and further work

This research study was requested by practitioners and has been adapted to their needs. We developed a new concept to monitor the evolution of a scar after the excision of a tumorous dermatosis.

This system is innovative because, within the same system, it offers hardware with an optical probe and software with image treatments along with a collaborative layer for telediagnosis.

In this paper we have shown the evolution of our optical probe which is original because it provides two

images at the same time and in the same position: functional with fluorescence and anatomic with white light. We described existing methods for image stitching, and then we proposed our implementation. Our approach is original in that we defined the best way to work on this kind of image (scar images), and then applied parameters obtained from stitching white light images on fluorescence images. Our collaborative layer allows practitioners to share images of scars, in real time, for a remote telediagnosis. Our choice to implement new components that adapt data in real time depending on the network and the terminal capacities of each user is also original.

With our implementations and tests, we show that the panoramic image process and adaptation presented in this paper allow our system to guarantee the best results possible. It is a reliable collaborative system, efficient enough to build panoramic scar images and to share them in an acceptable time frame.

Plans for future work are as follows:

1. The next step is to validate concepts through trials on *ex-vivo* tissues followed by a clinical study to completely validate the system in collaboration with several hospitals.
2. The fiber optical probe suffers from the fact that we need a mechanical device to move the fibers. We are currently developing a new kind of probe based on a CCD camera thus allowing a more simple and accurate imaging mode. The acquisition process time will also be improved by a factor of 20.
3. As we explained in Section 3.2. we are working on the possibility of integrating new components into the stitching step to take into account practitioners' additional needs.
4. At the collaborative work management layer, we are working on traceability. It is very important during a medical telediagnosis to trace the actions of practitioners. A new way to do this is to use ontology and to add a semantic on image transformations. Perhaps it would be interesting to use watermarking to keep information directly inside the images.

Acknowledgments

The authors thank the European Community (European FEDER) for financing a part of this work by the InterRegIV project SERVASTIC: Telediagnosis in Mobile Environments. The authors also thank the Franche-Comte Regional Council for financing some of our work through the Dicoderm Project.

References

- [1] S.A. de Araújo and H.Y. Kim, Ciratefi: An RST-invariant template matching with extension to color images, *Integrated Computer-Aided Engineering* **18**(1) (2011), 75–90.
- [2] J.-B. Aupet, N. Elmarzouqi, E. Garcia and J.-C. Lapayre, Virtual awareness card for adaptability in collaborative virtual environments, In *IEEE SETIT 2009*, Hammamet, Tunisia (March 2009), 228–234.
- [3] J.-B. Aupet, R. Kassab and J.-C. Lapayre, Wava: a new web service for automatic video data flow adaptation in heterogeneous collaborative environments, In *CDVE2009*, Luxembourg City, Springer-Verlag in LNCS, (September 2009), 280–288.
- [4] H. Bay, A. Ess, T. Tuytelaars and L. Van Gool, Speeded-up robust features (SURF), *Computer Vision and Image Understanding (CVIU)* **110**(3) (2008), 346–359.
- [5] V. Bonnans, T. Gharbi, C. Pieralli, B. Wacogne and P. Humbert, New fluorescence imaging probe with high spatial accuracy for *in vivo* applications, *Journal of Biomedical Optics* **9**(5) (2004), 928–933.
- [6] G. Bradski and A. Kaehler, Learning OpenCV, Book published by O'Reilly Media, Inc, sept 2008, ISBN 978-0-596-51613-0, 2001.
- [7] M. Brown and D.G. Lowe, Automatic panoramic image stitching using invariant features, *Int J of Computer Vision* **74**(1) (2007), 59–73.
- [8] F. Gouaux, Ambient intelligence and pervasive systems for the monitoring of citizens at cardiac risk: New solutions from the epi-medics project, In *Computers in Cardiology* **29** (2002), 289–292.
- [9] J. Hou, Z. Chen, X. Qin and D. Zhang, Automatic image search based on improved feature descriptors and decision tree, *Integrated Computer-Aided Engineering* **18**(2) (2011), 167–180.
- [10] P. Humbert, T. Gharbi, V. Bonnans, C. Pieralli and B. Wacogne, Fluorescence sensor, patent EP1300674, 2003.
- [11] S.-K. Hwang, M. Billingham and W.-Y. Kim, Local descriptor by zernike moments for real-time keypoint matching, in *IEEE CISP'08*, Washington, DC, USA **2** (2008), 781–785.
- [12] K. Jae-Gon et al., An optimal framework of video adaptation and its application to rate adaptation transcoding, *ETRI Journal* **27**(4) (2005), 341–354.
- [13] K.L. Kropfberger and H. Hellwagner, Quality variations of different priority-based temporal video adaptation algorithms, *IEEE Workshop on Multimedia Signal Processing* (2004), 183–186.
- [14] J. Lavanya et al., Distributed personal health information management system for dermatology at the homes for senior citizens, *C.B.; EMBS '06, 28th Annual International Conference of the IEEE Digital Object Identifier: 10.1109/IEMBS.2006.260760*, (2006), 6312–6315.
- [15] T. Lindeberg, Feature detection with automatic scale selection, In *Int J of Computer Visio* **30**(2) (1998), 79–116.
- [16] C. Louberry, M. Dalmau and P. Roose, Software architecture for dynamic adaptation of heterogeneous applications, In *Proceedings of the 8th ACM Int Conf on New Technologies in Distributed Systems*, Lyon, France (2008), 1–7.
- [17] D.G. Lowe, Distinctive image features from scale-invariant keypoints, *Int J of Comp Vision* **60**(2) (2004), 91–110.
- [18] D.G. Lowe, Object recognition from local scale-invariant features, *International Conference on Computer Vision*, Corfu, Greece (September 1999), 1150–1157.

- [19] C. Mailhes et al., The ursafe telemedicine project improving health care of the elderly, In *European Federation for Medical Informatics*, MIE, Saint Malo, France (May 2003).
- [20] J. Marescaux et al., Transatlantic robotassisted telesurgery, In *Nature* **413** (2001), 379–380.
- [21] J. Marescaux et al., Transcontinental robot assisted remote telesurgery: Feasibility and potential applications, In *Annal of Surgery*, Ann Surg **235** (2002), 487–492.
- [22] A.L.D. Martins, N.D.A. Mascarenhas and C.A.T. Suazo, Spatio-temporal resolution enhancement of vocal tract mri sequences based on image registration, *Integrated Computer-Aided Engineering* **18**(2) (2011), 143–156.
- [23] Le Mer et al., Argonaute 3d: a real-time cooperative medical planning software on dsl network, *12th Annual Medicine Meets Virtual Reality Conference*, Newport Beach, California **98** (January 2004), 203–209.
- [24] K. Mikolajczyk and C. Schmid, Indexing based on scale invariant interest points, In *ICCV'01* **1** (2001), 525–531.
- [25] J.-M. Morel and G. Yu, ASIFT: A new framework for fully affine invariant image comparison, *Intern Int IEEE Conf ICASSP'09* (April 2009), 1597–1600.
- [26] J.C. Noyer, P. Lanvin and M. Benjelloun, Correlation-based particle filter for 3D object tracking, *Integrated Computer-Aided Engineering* **16**(2) (2009), 165–177.
- [27] M. Ozuysal, P. Fua and V. Lepetit, Fast keypoint recognition in ten lines of code, In *Proc Int Conf on Computing Vision and Pattern Recognition*, IEEE (June 2007), 1–8.
- [28] L. Qin et al., Local invariant descriptor for image matching, In *Int IEEE Conf on Acoustics, Speech and Signal Processing* **2** (2005), 1025–1028.
- [29] P. Roose, M. Dalmau and F. Luthon, A distributed architecture for cooperative and adaptative multimedia applications, In *Proceedings of the 26th IEEE Int Computer Software and Applications Conf on Prolonging Software Life: Development and Redevelopment* (2002), 444–449.
- [30] C.N. Senarathne, S. Ransiri, P. Arangala, A. Balasooriya and C. De Silva, A faster image registration and stitching algorithm, *Industrial and Information Systems (ICIIS), 2011 6th IEEE Int Conf on Digital Object* (2011), 66–69.
- [31] P. Simard, L. Bottou, P. Haffner and Y. LeCun, Boxlets: A fast convolution algorithm for signal processing and neural networks, In *NIPS'98* (1998), 571–577.
- [32] S. Treuillet, B. Albouy and Y. Lucas, Three-dimensional assessment of skin wounds using a standard digital camera, *IEEE Transactions on Medical Imaging* (2009), 752–762.
- [33] P. Turcot and D.G. Lowe, Better matching with fewer features: The selection of useful features in large database recognition problems, In *ICCV Workshops* (2010), 2109–2116.
- [34] P.A. Viola and M.J. Jones, Rapid object detection using a boosted cascade of simple features, In *CVPR 2001* **1** (2001), 511–518.
- [35] X. Wang, F.C. Wu and Z. Wang, Harris feature vector descriptor (hfvd), In *ICPR'08* (2008), 1–4.
- [36] R. Watkins, E-learning tool for training and professional development services, e-learning, development of knowledge and/or skills for building competence, John Wiley & Sons, Inc. Hoboken, NJ, USA, Chapter 24, 2010.
- [37] S.A.J. Winder and M. Brown, Learning local image descriptors, In *CVPR* (2007), 1–8.
- [38] X. Yingen and K. Pulli, Sequential image stitching for mobile panoramas, *ICICS'09, Digital Object 2009* (2009), 1–5.
- [39] J. Ziegler et al., Context adaptive interaction for collaborative work, In *Proceedings of ACM CHI EA '10*, New York (2010), 4461–4464.